Bilbro, N., Hirst, A., Paez, A., Vasey, B., Pufulete, M. M., Sedrakyan, A., McCulloch, P. (2020). The IDEAL Reporting Guidelines: A Delphi Consensus Statement Stage specific recommendations for reporting the evaluation of surgical innovation. *Annals of Surgery*. https://doi.org/10.1097/SLA.0000000000004180

The IDEAL Reporting Guidelines: A Delphi Consensus Statement

Stage specific recommendations for reporting the evaluation of surgical innovation

Nicole Bilbro [a,b,*], Allison Hirst [a*], Arsenio Paez [c,d *], Baptiste Vasey [a], Maria Pufulete [e], Art Sedrakyan [f], Peter McCulloch [a]

*Joint first authors

Nicole A. Bilbro MD, MPH a,b

Allison Hirst MSc a

Arsenio Paez PT, DPT c,d

Baptiste Vasey, a

Maria Pufulete, e

Art Sedrakyan MD, PhD f

Peter McCulloch MD a

On behalf of the IDEAL Collaboration Reporting Guidelines Working Group

a Oxford University Hospitals

IDEAL Collaboration

Nuffield Department of Surgery

John Radcliffe Hospital

Oxford, UK

b Maimonides Medical Center

Department of Surgery

Brooklyn, NY, USA

c Nuffield Department of Primary Care Health Sciences

University of Oxford

Oxford, UK

d Northeastern University

Department of Physical Therapy, Movement, and Rehabilitation Sciences

Boston, Massachusetts, USA

e Bristol details

f Weill Cornell Medical College

New York, NY, USA

Corresponding Author:

Peter McCulloch

Professor of Surgical Science & Practice

University of Oxford

Nuffield Department of Surgical Sciences (Level 6)

John Radcliffe Hospital

Headington

Oxford OX3 9DU

Tel. +44 (0)1865 612276

Email. peter.mcculloch@nds.ox.ac.uk

Funding: No specific funding was obtained for this project. The IDEAL Collaboration is supported by the Oxford NIHR Biomedical Research Centre, an unrestricted Educational Grant from Medtronic as well as other educational grants and meetings income. (AH and PM).

Running Head: IDEAL Reporting Guidelines Statement

ABSTRACT

Objective: To define reporting standards for IDEAL format studies.

Background: The IDEAL Framework and Recommendations establish an integrated pathway for evaluation of new surgical techniques and complex therapeutic technologies. However guidance on implementation has been incomplete, and incorrect use is commonly seen. We describe the consensus development of reporting guidelines for the IDEAL stages, and plans for their dissemination and evaluation.

Methods: Using the EQUATOR Network recommendations, participants with knowledge of IDEAL were surveyed to determine which IDEAL stages needed reporting guidelines. Draft checklists for stages 1,2a, 2b and 4 were subsequently developed by 3 researchers (NB, AH, PMcC), and revised through a two-round Delphi consensus process. A final consensus teleconference resolved outstanding disagreements and clarified wording for checklist items.

Results: 61 participants completed the initial survey, a clear majority indicating that new reporting guidelines were needed for IDEAL Stage 1 (69.5%), Stage 2a (78%), Stage 2b (74.6%), and Stage 4 (66%). A proposed set of checklists was modified by survey participants in two online Delphi rounds (n=54 and n=47 respectively), resulting in a penultimate checklist for each stage. Fourteen expert working group members finalised the checklist items and successfully resolved any outstanding areas without agreement on a consensus call.

Conclusions: Participants familiar with IDEAL called for reporting guidelines for studies in all IDEAL stages except stage 3. The checklists developed have the potential to improve standards of reporting and thereby advance the quality of research on surgery and complex interventions and technologies, but require further evaluation in use.

Introduction

Surgical innovation is recognized as having unique methodological and practical barriers that make rigorous evaluation and quality reporting inherently complex.[1] The IDEAL (Idea, Development, Exploration, Assessment, Long Term Study) Framework and Recommendations were introduced in 2009 to clarify and address the key challenges presented in surgical innovation [2]. The IDEAL integrated evaluation pathway guides the conduct and reporting of clinical studies of the outcomes of surgery and complex therapeutic interventions. The IDEAL Framework has evolved over time, and now includes a model for medical devices evaluation, IDEAL-D, and a modified framework for physical therapy and rehabilitation, IDEAL-Physio[3,4,5,6].

Although widely accepted in principle, the use of the IDEAL recommendations in practice to date has been limited and in many cases sub-optimal. Surveys of publications using IDEAL have shown that reporting of the stages was often inaccurate and the key features which define each stage were frequently omitted[7,8,9]. The IDEAL Recommendations introduce several novel ideas into clinical study design, and following other published guidelines alone is therefore unlikely to enable authors to report them accurately. The lack of clear and detailed instructions on reporting has been considered to be a barrier to wider adoption of IDEAL, and there have been calls for the development of reporting guidelines specific to IDEAL [10,11].

Reporting Guidelines

The EQUATOR Network (http://www.equator-network.org) defines a reporting guideline as a simple, structured tool for health researchers to use while writing manuscripts to guide the authors in reporting a specific type of research[12]. Guidelines are developed using an

explicit methodology and provide a minimum list of information needed to ensure a manuscript can be understood and assessed by a reader, replicated by a researcher, used to make a clinical decision, and included in a systematic review. There are now over 400 checklists on the EQUATOR Network's website, including 42 that are surgery related (http://www.equator-network.org/reporting-guidelines/). Among these existing reporting guidelines applicable to surgery SCARE [13], PROCESS [14], STROCSS [15] and CONSORT NPT [16] aim to improve the reporting of specific, well-recognised study types (case reports, case series, cohort studies and randomized controlled trials). Additional guidance exists for specific aspects of surgical evaluative study designs. For example TIDIER [17] has been developed as an extension of CONSORT to improve reporting of intervention details, and the COHESIVE [18] guidelines are being developed for a generic core outcome set in studies of early phase surgery and medical devices. A CONSORT extension for RCTs using cohorts and routinely collected health data is also underway [19]. Additionally the National Evaluation System for Health Technology Coordinating Center (NESTcc) Methods Framework has been developed to define the key components of study design, protocols and evidence gathering for the evaluation of medical devices [20]

All these guidelines draw upon principles of good research methodology, as does the IDEAL framework. However the IDEAL reporting guidelines have a unique and specific focus on stepwise evaluation of surgical innovation through all stages of its development, from pre-clinical to first in human evidence, feasibility and randomized trials to long-term follow-up and real-world data. Throughout these stages, IDEAL prescribes specific novel study formats, focussing on clarity about key elements of surgical innovation with which many researchers may be unfamiliar. This likely explains the inaccuracy in reporting noted in studies

attempting to follow the IDEAL recommendations and suggests a pressing need for IDEAL reporting guidelines.

The development of such IDEAL specific guidelines would serve to improve the usability and impact of IDEAL. As Altman and Moher state, "reporting problems affect journal articles in two main ways. First, the study methods are frequently not described in adequate detail. Second, the study findings are presented ambiguously, incompletely, or selectively. The cumulative effect of these problems is to render many reports of research unusable or even harmful; at the very least, such papers certainly represent a waste of resources"[21]. Chalmers and Glasziou have highlighted the detrimental effects of waste in the production and reporting of research evidence and call for increased use of reporting guidelines as part of the solution. [22]

This paper describes the development of a set of IDEAL reporting guidelines which aim to improve the quality, transparency, consistency, and utility of the surgical innovation evidence base. The protocol for the project was published in 2018. [23]

Methods

Research Design:

We developed the IDEAL reporting guidelines through a five-phase process between October 2018 and May 2019. This comprised (1) identifying existing relevant reporting guidelines and those in development registered on the EQUATOR Network website (http://www.equator-network.org), (2) a survey to establish which stages of IDEAL, if any, participants considered in need of new reporting guidelines, (3) the development of draft checklists for Stages 1, 2a, 2b and 4 following establishment of need based on consensus, (4)

two rounds of a Delphi consensus survey to refine each of the checklists, and (5) a consensus meeting via teleconference with the working group to finalise the checklists.

Participants:

Using our MailChimp e-newsletter, we contacted 445 individuals who met at least one of four criteria: (1) member of the original IDEAL steering group; (2) a wider member of the IDEAL Collaboration signed up to receive newsletters; (3) author or co-author of any key core IDEAL publication; (4) attended one or more of the 2016, 2017, and 2018 IDEAL conferences. From this group, 119 expressed interest in participating, and those individuals were invited to complete an online survey about the need for reporting guidelines for any stage of IDEAL, as defined by the EQUATOR Network's suggested criteria. [24] The 61 respondents recommended that new reporting guidelines should be developed for Stage 1 (69%), Stage 2a (78%), Stage 2b (75%), and Stage 4 (66%) (***Supplementary Figure 1***). IDEAL Stage 3 was considered adequately covered by existing guidelines (CONSORT, CONSORT NPT and other relevant extensions) by 68% of respondents. Characteristics of responding participants are shown in ***Table 1***. Fourteen volunteers from groups (1) and (3) formed a working group to agree the final checklists during the final consensus call.

Development of Checklists:

The initial drafts of checklists for the four stages were developed using a combination of approaches to identify what key items may be needed. These included ongoing monitoring of the literature using or commenting on the usability of IDEAL identified from Google Scholar and Web of Science Alerts of papers citing the 2009 and 2019 IDEAL Statement papers. Assessing gaps in existing reporting guidelines, and accumulated refinements of IDEAL principles developed from issues repeatedly raised in recent years during educational

seminars and lectures on IDEAL. These checklists were evaluated in two rounds of a Delphi consensus survey using a nine-point Likert scale, (as recommended by GRADE), to rate agreement on the importance of each element and elicit free-text comments on the wording of each item[25,26]. In the Likert scale, a score of 1–3 signified limited importance of the item, 4–6 important but not critical, and 7–9 critical. If 70% or more of respondents scored an item 7–9 and fewer than 15% scored it 1–3, then that item was included in the reporting guideline. Conversely, consensus that an item should not be included was defined as 70% or more scoring it 1–3 and 15% or less scoring it 7–9. Following round one, the authors reviewed the responses (n=54) and two reviewers (NB, AH) modified the checklist drafts based on the participants scoring and feedback, editing phrasing and wording of checklist items and merging redundant items, but ensuring meaning was not significantly changed. Round two of the Delphi survey presented the results of round one to participants, addressed clarifications and served to refine the checklist items where consensus was not yet reached. Participants were again asked to score the Round 2 checklist items on a 1-9 Likert Scale.

The responses from round two (n=47) were compiled for review by 2 researchers (NB, AH) and presented to the working group during the final teleconference consensus meeting.

Consensus Teleconference:

A consensus call held in May 2019 was attended by eleven of the fourteen members of the working group. Three participants who were unable to attend submitted written contributions via email. A draft of the methods, checklist drafts for each stage, and a provisional list of items where the Delphi process had not yielded a definitive result were prepared in advance and sent to the call participants. The items that remained controversial

were discussed and modifications were made once the group reached agreement. All checklist items for each stage were then reviewed and revised as needed to ensure clarity. The group also decided on a strategy to disseminate the checklists. The revisions were then sent out to the participants for final comment and approval.

Results

The citation alerts and search of the EQUATOR Network list of reporting guidelines revealed no specific guidelines dealing with the IDEAL Recommendations as an entity either as the sole focus of study or as part of a larger ambit, and no guidelines which specifically described how to report IDEAL format studies. Numerous guidelines contained generic recommendations also included by IDEAL, e.g. use of standard outcome definitions, publication of protocols and full informed consent, but none referenced IDEAL as a special case. [14-19]

Supplementary Figures 2 and 3 show the distribution of Scores for each checklist item in Rounds 1 and 2 respectively of the Delphi process. The 4 unresolved issues following Round 2, discussed during the final consensus call were: Inclusion of IDEAL and the stage number in titles; inclusion of "first-in-human" in the title of Stage 1 studies; mandatory publication of a prior protocol for all studies; and how to resolve tensions between commercial confidentiality and ethical demands for transparency in early stage studies involving devices. The solutions to each are incorporated in the final full checklists presented here.

IDEAL Checklists

IDEAL is based on ethical and methodological principles which apply to all stages and types of medical research, and many of which are held in common with other methodological

paradigms and checklists. ***Table 2*** presents the full checklists for IDEAL Stages 1, 2a, 2b and 4 and highlights key or unique IDEAL items with an example from the literature which illustrates appropriate use. The quoted paper does not necessarily demonstrate good reporting for other IDEAL recommendations.

Discussion

The IDEAL Recommendations include several innovations in study design for clinical research, such as reporting of iterative modifications and monitoring of learning curves. Because of this, the diligent use of other guidelines cannot help investigators to report IDEAL studies appropriately, and an IDEAL-specific guideline is required. For example comparing the SCARE guidelines for surgical case reports [13] and the IDEAL Stage 1 recommendations, IDEAL specifies requirements absent from SCARE (Applicability only to first-in-human studies, justify need and refer to pre-clinical development, detailed account of patient selection, requirement for fully informed consent for treatment, requirement for explanation of procedure detailed enough to allow reproduction).

The IDEAL checklists are intended to provide a minimum list of concepts authors should include in a report of an IDEAL format study. They have been created for wide-spread adoption across all interventional specialties, including therapeutic devices, for which IDEAL has published a modified framework, IDEAL-D[5]. The checklists are equally usable for IDEAL and IDEAL-D studies, most items pertaining to both surgical and device innovation. While not designed expressly for these purposes, the guidelines can also be used both prospectively to help plan a study and retrospectively to assist in appraisal. They can also be used alongside other reporting guidelines to provide additional useful detail (e.g. COHESIVE

for core outcomes reporting, CONSORT extension for routinely collected health data for Stage 4). [18,19]

Dissemination and evaluation of IDEAL checklists

Checklists are not effective unless used, so dissemination efforts are important. The EQUATOR network [27] recommends 10 specific actions to aid dissemination: making the checklist freely accessible in editable format; publishing and presenting editorials, blogs, website pages, conference papers and webinars, co-operating with EQUATOR, seeking journal endorsements and running training courses. We intend to follow this guidance utilising our website and Twitter accounts (www.ideal-collaboration.net and @IDEALCollab). Some specific elements of our strategy will be the use of social media to establish a feedback dialogue with users, presentation of the guidelines at conferences and scientific meetings, editorials and blogs, and translation into Chinese and Spanish to maximise international readership. In line with EQUATOR guidance, we will pilot the checklist in ongoing IDEAL format studies, for which we will shortly be introducing an advisory service, and will publish an updated version with any modifications resulting from both feedback and piloting. A pre-post guideline publication survey of reporting of IDEAL format studies to evaluate the effectiveness of these efforts will be developed at a future date. Endorsement by authoritative bodies and efforts by editors to encourage or enforce compliance were features of the development and spread of the CONSORT guidelines, which appear to be the best complied with to date, so we will continue to seek support in these sectors.

Limitations

Like all consensus processes, our Delphi study was dependent on the representativeness of participants. We invited participants with a good understanding of IDEAL, as a naïve group

would be unlikely to be able to appreciate some of the issues or achieve consensus. We attempted to include a wide range of stakeholders but were not universally successful. The number of journal editors, for example, was small. A different group might have come up with different recommendations, but no major differences in responses could be detected according to background in our sample. Although consensus was achieved by GRADE criteria on the need for a guideline in all stages except Stage 3, the minority which disagreed was substantial (around 25%) in each case. Informal feedback indicates that this may have been because participants familiar with IDEAL had come to see reporting of IDEAL stage items as straightforward, and not needing guidance. Some participants also mentioned the proliferation of reporting guidelines as a reason for caution.

Conclusions

The consensus-based checklists developed through this Delphi process have the potential to improve the standards of reporting of early stage innovation and surgical trials and thereby advance the quality of research on surgery and complex interventions. Surgical innovation and research that follows a transparent, sequential framework such as IDEAL and is reported with these guidelines can also aid reproducibility, reduce research waste and improve evidence-based practice for the benefit of patients.

Contributors

IDEAL Collaboration Reporting Guidelines Working Group: Jane Blazeby, Maroeska Rovers, Christopher Pennell, Joel Horovitz, Angelos Kolias, Janet Martin, Tom Lewis, Riaz Agha, Josh Feinberg, in addition to the authors.

Acknowledgments

Delphi Survey Participants: Alfred Adiamah, Augusto Azuara-Blanco, Chris Bainbridge, Dimitri Barski, Natalie Blencowe, Jan Bosteels, Cristina Camargo, Wim Ceelen, Alejandro Correa Paris, Neil Corrigan, Joanna Crocker, Nathan Curtis, Prita Daliya, Susan Dutton, Thorsten Ecke, Julio F Fiore Jr., George Garas, Herney Garcia-Perdomo, Gus Gazzard, Janneke Grutters, Emma Hall, Thomas Hamilton, Ted Heise, Daniel Hind, Sorin Hostiuc, Jarkko Jokihaara, Emma Jones, Lileswar Kaman, Sean Kehoe, Samuel Lawday, Aaron Lawson McLean, Daniel Liu, Mario Lobao, Robert Michael Lundin, Karen MacPherson, Guy Maddern, Varaidzo Mapunde, John Mathews, Jose G Merino, Andrew Metcalfe, Richard Naspro, Ionut Negoi, Stefano Olgiati, Thomas Pidgeon, Pascal Probst, Sarah Prowse, Wendy Rogers, Camiel Rosman, Laura Rossi, Suzanne Ryan, Franziska Saxer, Markus Schafer, Patrick Sequeira-Byron, Neil Smart, Diederik Smeeing, Matthew Smith, Grant Stewart, William Summerskill, Olga Tucker, Emma Upchurch, Bolivar Alfonso Villacis Bermeo, Gustavo Zanoli

References

1. Ergina PL, Cook JA, Blazeby JM, et al. Challenges in evaluating surgical innovation. *Lancet*. 2009;374(9695):1097-1104.

2. McCulloch P, Altman DG, Campbell WB, et al. No surgical innovation without evaluation: the IDEAL recommendations. *Lancet*. 2009;374(9695):1105-1112.

3. Pennell CP, Hirst AD, Campbell WB, et al. Practical guide to the Idea, Development and Exploration stages of the IDEAL Framework and Recommendations. *Br J Surg*. 2016. Apr;103(5):607-615.

4. Hirst A, Philippou Y, Blazeby J, et al. No Surgical Innovation Without Evaluation: Evolution and Further Development of the IDEAL Framework and Recommendations. *Ann Surg*.2019 Feb;269(2):211-220.

5. Sedrakyan A, Campbell B, Merino JG, et al. IDEAL-D: A rational framework for evaluating and regulating the use of medical devices. *BMJ*. 2016;353(June):1-7.

6. Hamilton D, Beard D, McCulloch P, et al. Evidence-Based Evaluation of Practice and Innovation in Rehabilitation Using the IDEAL-Physio Framework. *Phys Ther.* 2018 Feb 1;98(2):108-121.

7. Khachane A, Philippou Y, Hirst A, et al. Appraising the uptake and use of the IDEAL Framework and Recommendations: A review of the literature. *Int J Surg*. 2018;57(July):84-90.

8. Tradewell MB, Albersheim J, Dahm P. Use of the IDEAL framework in the urological literature: where are we in 2018? *BJU Int*. 2019:1078-1085.

9. Currie A, Brigic A, Blencowe NS et al. Systematic review of surgical innovation reporting in laparoendoscopic colonic polyp resection. Br J Surg. 2015 Jan;102(2):e108-116.

10. Barkun JS, Dimick JB, Clavien PA. Surgical Research in Patients: Ideal Time for an IDEAL Checklist. *Ann Surg*. 2019;269(2):208-210.

11. Kirkham EN, Main BG, Jones KJB, et al. Systematic review of the introduction and evaluation of magnetic augmentation of the lower oesophageal sphincter for gastro-oesophageal reflux disease. *Br J Surg*. 2020;107(1):44–55.

12. EQUATOR Network website. What is a Reporting Guideline? Available at: https://www.equator-network.org/about-us/what-is-a-reporting-guideline/ (Accessed 30th March 2020)

13. Agha RA, Borrelli MR, Farwana R et al. For the SCARE Group. The SCARE 2018 Statement: Updating Consensus Surgical CAse REport (SCARE) Guidelines, *Int J Surg.* 2018;60:132-136

14. Agha RA, Borrelli MR, Farwana R et al. For the PROCESS Group. The PROCESS 2018 Statement: Updating Consensus Preferred Reporting of CasE Series in Surgery (PROCESS) Guidelines, *Intl J Surg.* 2018;60:279-282

15. Agha R, Abdall-Razak A, Crossley E et al. STROCSS 2019 Guideline: Strengthening the reporting of cohort studies in surgery. *Int J Surg*. 2019;72:156-165.

16. Boutron I, Altman DG, Moher D, et al; CONSORT NPT Group. CONSORT Statement for Randomized Trials of Nonpharmacologic Treatments: A 2017 Update and a CONSORT Extension for Nonpharmacologic Trial Abstracts. *Ann Intern Med*. 2017;167(1):40-47

17. Hoffmann T, Glasziou P, Boutron I, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ.* 2014;348:g1687

18. Avery K, Blazeby J, Wilson N, et al. Development of reporting guidance and core outcome sets for seamless, standardised evaluation of innovative surgical procedures and devices: a study protocol for content generation and a Delphi consensus process (COHESIVE study). BMJ Open. 2019 Sep 12;9(9):e029574.

19. Kwakkenbos L, Juszczak E, Hemkens LG, et al. Protocol for the development of a CONSORT extension for RCTs using cohorts and routinely collected health data. *Res Integr Peer Rev*. 2018;3(1):1-9. doi:10.1186/s41073-018-0053-3

20. National Evaluation System for health Technology Coordinating Center (NESTcc) Frameworks Feb 2020. https://mdic.org/resource/nestcc-data-quality-framework/ and https://mdic.org/resource/nestcc-methods-framework/ (Accessed 30th March 2020)

21. Altman DG, Moher D. *Guidelines for Reporting Health Research: A User's Manual*. (Moher D, Altman DG, Schulz KF, Simera I, Wagner E, eds.). John Wiley & Sons, Ltd; 2014. https://onlinelibrary.wiley.com/doi/book/10.1002/9781118715598.

22. Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. Lancet. 2009 Jul 4;374(9683):86-89.

23. Agha RA, Hirst A, Khachane A, et al. A protocol for the development of reporting guidelines for IDEAL stage studies. *Int J Surg Protoc*. 2018;9:11-14.

24. EQUATOR Network - Identifying the need for a new reporting guideline. Available at: http://www.equator-network.org/toolkits/developing-a-reporting-guideline/identifying-the-need-for-a-new-reporting-guideline/ (Accessed 30 March 2020).

25. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. 2008;336(7650):924 LP - 926.

26. Jaeschke R, Guyatt GH, Dellinger P, et al. Use of GRADE grid to reach decisions on clinical practice guidelines when consensus is elusive. *BMJ*. 2008;337. doi:10.1136/bmj.a744

27. EQUATOR Network – Developing and disseminating your reporting guideline. Available at: https://www.equator-network.org/toolkits/developing-a-reporting-guideline/disseminating-your-reporting-guideline/ (Accessed 30 March 2020)